Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 5 (2015) 66 – 81

Transportation Research Procedia

www.elsevier.com/locate/procedia

SIDT Scientific Seminar 2013

Using the General Link Transmission Model in a Dynamic Traffic Assignment to simulate congestion on urban networks

Guido Gentile *

Dipartimento di Ingegneria Civile Edile Ambientale, Università di Roma "La Sapienza", Via Eudossiana 18 – 00184 Rome, Italy

Abstract

This article presents two new models of Dynamic User Equilibrium that are particularly suited for ITS applications, where the evolution of vehicle flows and travel times must be simulated on large road networks, possibly in real-time. The key feature of the proposed models is the detail representation of the main congestion phenomena occurring at nodes of urban networks, such as vehicle queues and their spillback, as well as flow conflicts in mergins and diversions. Compared to the simple word of static assignment, where only the congestion along the arc is typically reproduced through a separable relation between vehicle flow and travel time, this type of DTA models are much more complex, as the above relation becomes non-separable, both in time and space.

Traffic simulation is here attained through a macroscopic flow model, that extends the theory of kinematic waves to urban networks and non-linear fundamental diagrams: the General Link Transmission Model. The sub-models of the GLTM, namely the Node Intersection Model, the Forward Propagation Model of vehicles and the Backward Propagation Model of spaces, can be combined in two different ways to produce arc travel times starting from turn flows. The first approach is to consider short time intervals of a few seconds and process all nodes for each temporal layer in chronological order. The second approach allows to consider long time intervals of a few minutes and for each sub-model requires to process the whole temporal profile of involved variables. The two resulting DTA models are here analyzed and compared with the aim of identifying their possible use cases.

A rigorous mathematical formulation is out of the scope of this paper, as well as a detailed explanation of the solution algorithm.

The dynamic equilibrium is anyhow sought through a new method based on Gradient Projection, which is capable to solve both proposed models with any desired precision in a reasonable number of iterations. Its fast convergence is essential to show that the two proposed models for network congestion actually converge at equilibrium to nearly identical solutions in terms of arc flows and travel times, despite their two diametrical approaches wrt the dynamic nature of the problem, as shown in the numerical tests presented here.

Peer-review under responsibility of the Società Italiana dei Docenti di Trasporti (SIDT).

Keywords: Dynamic User Equilibrium; Dynamic Network Loading; Theory of Kinematic Wave; time discretization; algorithm convercence; macroscopic flow models, queue spillback, Intelligent Transport Systems, traffic forecast.

* Corresponding author. Tel.: +39-06-44585-737; fax: +39-06-44585-129.
E-mail address: guido.gentile@uniroma1.it

Peer-review under responsibility of the Società Italiana dei Docenti di Trasporti (SIDT).
doi:10.1016/j.trpro.2015.01.011

1. Background and motivation

Despite 20 years of intensive efforts produced by the transportation research community (starting from the early models of, e.g., Jayakrisham et al., 1994; Ben-Akiva et al., 1997; Adamo et al., 1999), Dynamic Traffic Assignment (DTA) is still one of the most challenging issues in network modelling. A satisfactory mathematical framework that ensures the existence and uniqueness of a dynamic user equilibrium, along with a convergent algorithm that can rapidly compute its solution, are goals to be yet achieved. Some consistent formulations are available (e.g., Friesz et al., 1993; Ran and Boyce, 1994; Heydecker and Addison, 1998); but one shall be ready to pay a high price in terms of model realism on the supply side (e.g. simple point-queue models with no flow conflicts at nodes and no spillback congestion) and/or on the demand side (e.g. instantaneous shortest paths, system optimum vs. user equilibrium). As a result, although new researches (e.g., Mahut and Florian, 2008; Smith and Mounce, 2011; Ramadurai and Ukkusuri, 2011) are trying to exploit some of the ideas that allowed a relevant improvement in the state-of-the-practice for static assignment (e.g. Dial, 2006; Gentile, 2014), fixed-point-like formulations, with their poor mathematical properties, and MSA (Method of Successive Averages) algorithms, with their slow convergence patterns, still seem to dominate the panorama of DTA models today (an up-to-date review of some academic and commercial software available for real-life applications can be found in Barcelo, 2010).

On the other hand, the need of reliable traffic models that can be efficiently solved also for large networks, is constantly increasing. Once the limited realism of the static approach is recognized (e.g. Gentile, 2010a), DTA is seen by more and more experts as a fundamental tool to analyze congested networks in two distinct use cases: off-line transport planning, to provide simulations of design and management scenario; real-time traffic monitoring, to provide short term predictions on flows and speeds. Traffic management shares elements from both base cases.

The above two applications of assignment models have a key difference. In traffic monitoring, the main aim of simulations is to reproduce vehicle flows and travel times on the road network for given demand and supply, possibly in real-time. In transport planning, the main aim of simulations is to compare different design scenarios. This diverse perspective implies that, while in the latter case it can be important to compute precise equilibrium solutions, even at the price of adopting a rough representation of traffic phenomena, such as that provided by static assignment models, the contrary is true in the former case.

Actually in the context of traffic monitoring the paradigm of equilibrium path choices is not as stringent (is there a dynamic equilibrium in real-time?), while we need supply models capable to reproduce the essence of network congestion, that is the evolution in time of queues and their possible spillback.

In other words, although we still consider the user equilibrium paradigm a convenient modelling framework, in traffic monitoring more importance is given to the realism of dynamic network loading (consistency of flows, travel times and capacity constraints) including vehicle interactions at nodes, than to the possibility of reaching a highly precise consistency between path choices and arc performances (i.e. times and costs). Then, in traffic management applications the essential property of the DTA model should be its sensitivity (sound reaction) to local changes of the main supply parameters, such as capacity reductions due to traffic events. To cope with this need, the dynamic models applied in real-time (e.g., Mahmassani, 2001; Ben-Akiva et al., 2002; Gentile and Meschini, 2011) introduce some simplifying assumption, especially wrt route choice, compared to static off-line models.

The main aim of this paper is to present two new models for DTA where the network congestion is reproduced through the General Link Transmission Model (Gentile, 2010b). The GLTM is a macroscopic flow model, that extends the theory of kinematic waves to urban networks and allows for any concave fundamental diagram. Like the Link Transmission Model (Yperman, 2007), the GLTM does not require to split each arc into smaller segments, as in the Cell Transmission Model (Daganzo, 1994-1995). Unlike the LTM, the GLTM propagates the demand flows from the origins through the network, based on given (possibly time varying) splitting rates (or turning fractions), without the need of handling destination specific flows; clearly, a given flow rate is also diverted from the network nodes to the destinations, and thus exits the network. However, this simplified approach, which allows to simulate large networks in real time, requires an iterative process to obtain a Dynamic Network Loading that is consistent with given Origin Destination flows. Indeed, even if the splitting rates are derived from the assignment to the network of the given OD flows, new travel times computed by the GLTM would produce a different result if flows would be propagated according to the same arc conditional probabilities toward the destination (i.e. no change in the

route choice is considered). Finally, unlike both LTM and CTM, the GLTM allows for the simulation of hypocritical congestion along the link.

The sub-models of the GLTM, namely the Node Intersection Model, the Forward Propagation Model and the Backward Propagation Model, can be combined in two different ways to produce arc travel times starting from turn flows. The first approach is to consider short time intervals of a few seconds and process all nodes for each temporal layer in chronological order; this is the classical way of dealing with the network simulation in DTA macro and meso-scopic models (see also, Papageorgiou, 1990; Lo and Stezo, 2012). The second approach allows to consider long time intervals of a few minutes and for each sub-model requires to process the whole temporal profile of involved variables; this innovative method has been first proposed in Bellei et al. (2005), and other researchers are now further developing it (e.g. Himpe et al., 2013). The two resulting models will be analyzed and compared with the aim identifying their possible use cases.

The DTA models proposed here develop from the deterministic version of the model presented in Bellei et al. (2005) for the case of Logit route choice, which has been implemented in the VISUM software package as the reference macroscopic method for dynamic assignment since version 10 (Gentile et al., 2006) and simply called DUE (Dynamic User Equilibrium).

In this paper, the dynamic equilibrium is sought through a new method based on Gradient Projection, which is capable to solve both the proposed DTA models with any desired precision in a reasonable number of iterations. Although the full presentation of the algorithm is out of the scope of this paper, its fast convergence was essential to show that the two distinct models for network congestion actually provide at equilibrium nearly identical solutions in terms of arc flows and travel times, despite their diametrical approaches to the dynamic nature of the problem. Numerical results prove the validity of this assertion for both dynamic equilibrium and network loading wrt different time discretizations, with time intervals of 6, 60 and 600 sec.

2. Conceptual framework

In this section the proposed DTA models are presented through a conceptual schema, as a fixed point problem. A more convenient mathematical formulation as a variational inequality problem with full explanation of all variables is out of the scope of this paper, but can has been recently presented in Gentile (2014b). Here we concentrate our attention on the essential aspects of DTA modelling architecture.

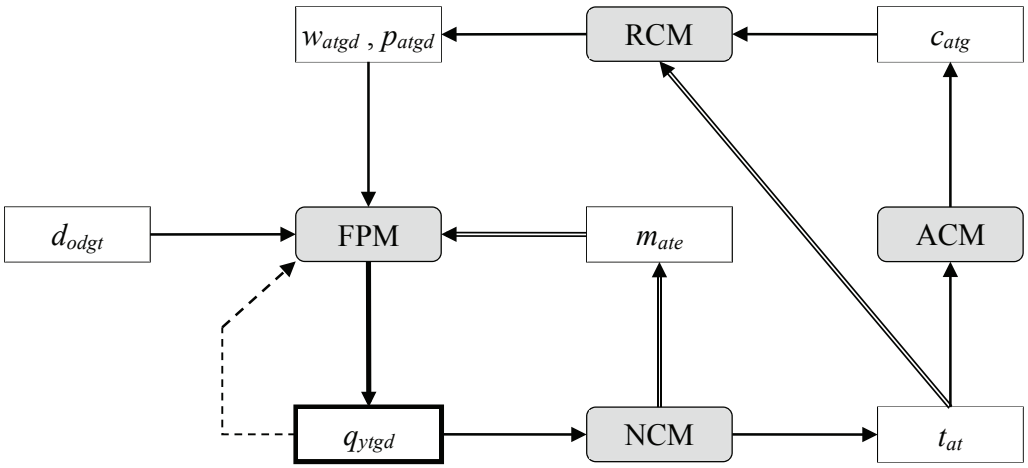

Figure 1. Schema of the DTA model and algorithm. The dashed arrow indicates that the iterate variable is used in the next iteration. The thick arrow recalls that an algorithmic transformation of the iterate (e.g. through MSA or Gradient Projection) is required to provide a descent direction for the convergence of the equilibrium as a fixed-point problem.

NCM – Network Congestion Model. It takes as input the turn flows, aggregated for all destinations, and yields as output the arc travel times. This model aims at reproducing different traffic phenomena, from hypocritical congestion, to queue spillback.

ACM – Arc Cost Model. It takes as input the travel times and yields the arc costs perceived by each user class, considering their different values of time and tolls.

RCM – Route Choice Model. It takes as input the arc costs, as well as the arc travel times that allow for the dynamic (backward) concatenation of perceived utilities. It yields as output the expected costs to reach the destination using each local alternative, that are then used to compute the arc conditional probabilities. However, it can be convenient to perform the latter computation directly in the FPM.

FPM – Flow Propagation Model. It takes as input the travel demand and the local choices, as well as the travel times (or the propagation map) that allow for the dynamic (forward) propagation of users. It yields as output the arc flows directed towards each destination.

Notation of the Dynamic User equilibrium

q_{ytgd}	flow of class $g \in G$ passing through turn $y \in Y$ during time interval $[\tau_t , \tau_{t+1})$ with $t \in T$ and directed toward destination $d \in Z$
m_{aet}	share of flow entering arc $a \in A$ during time interval $[\tau_t , \tau_{t+1})$ with $t \in T$ and exiting during time interval $[\tau_e , \tau_{e+1})$ with $e \geq t \in T$
t_{at}	travel time of arc $a \in A$ for users entering at time τ_t , with $t \in T$
c_{atg}	cost of arc $a \in A$ perceived by users of class $g \in G$ entering at time τ_t , with $t \in T$
w_{atgd}	expected cost perceived by users of class $g \in G$ users entering arc $a \in A$ during time interval $[\tau_t , \tau_{t+1})$ with $t \in T$ and directed toward destination $d \in Z$
p_{atgd}	probability that, during time interval $[\tau_t , \tau_{t+1})$ with $t \in T$, users of class $g \in G$ directed toward destination $d \in Z$ choose to enter arc $a \in A$ conditional on being at its initial node
d_{odgt}	demand flow of class $g \in G$ travelling from origin $o \in Z$ to destination $d \in Z$ and leaving during time interval $[\tau_t , \tau_{t+1})$ with $t \in T$

For the case of deterministic choices we consider the temporal layer computation of dynamic shortest paths proposed in Gentile et al. (2004), which extends to continuous temporal profiles the approach of Chabini (1998), instead of the trajectory based computation proposed by Pallottino and Scutellà (1998). Consistently with the time discretization, we consider flows as constant during each time interval. Then, all users leaving a node during a same interval choose the next arc toward their destination in the same way and shall evaluate the different local alternatives considering the expected cost of the last instant; this way, all such users will be affected by their own choice. On the contrary case, the model becomes instable.

Two kind of arc flows will be considered in the model. The first one, produced by the Flow Propagation Model on the basis of route choices to determine splitting rates, is consistent with drivers' behaviour but inconsistent with congestion phenomena. The second one, produced by the Network Congestion Model for given splitting rates (see later), is consistent with congestion phenomena but inconsistent with drivers' behaviour. Only at equilibrium these two flow variables become mutually consistent.

In this paper, we investigate two different approaches for the NCM, that are based on the same components and describe the same phenomena, but diverge substantially in terms of algorithm implementation and granularity of time discretization.

2.1. The General Link Transmission Model and the simulation approach

The first approach is based on the General Link Transition Model (Gentile, 2010b) as a simulation tool. This is a macroscopic model based on the solution in terms of cumulative flows of the kinematic wave model (Newell, 1993) suitably extended to allow any concave fundamental diagram. For each time interval in chronological order, all nodes are processed in no particular order. Such an approach based on temporal layers allows for the exact solution

of the model, under the assumption that the speed of no wave, hypocritical (free flow speed) or hypercritical (jam wave speed), is faster than the simulation speed (the ratio between the length of the arc and the duration of the time interval). This way, no arc inflow or outflow in a given time interval can affect the flow state of other arcs in the same time interval. In practice, we can stretch le length of some very short link so as to consider a simulation tick of up to 6-12 seconds; otherwise, we can stick to a classical time discretization with intervals of 1 second.

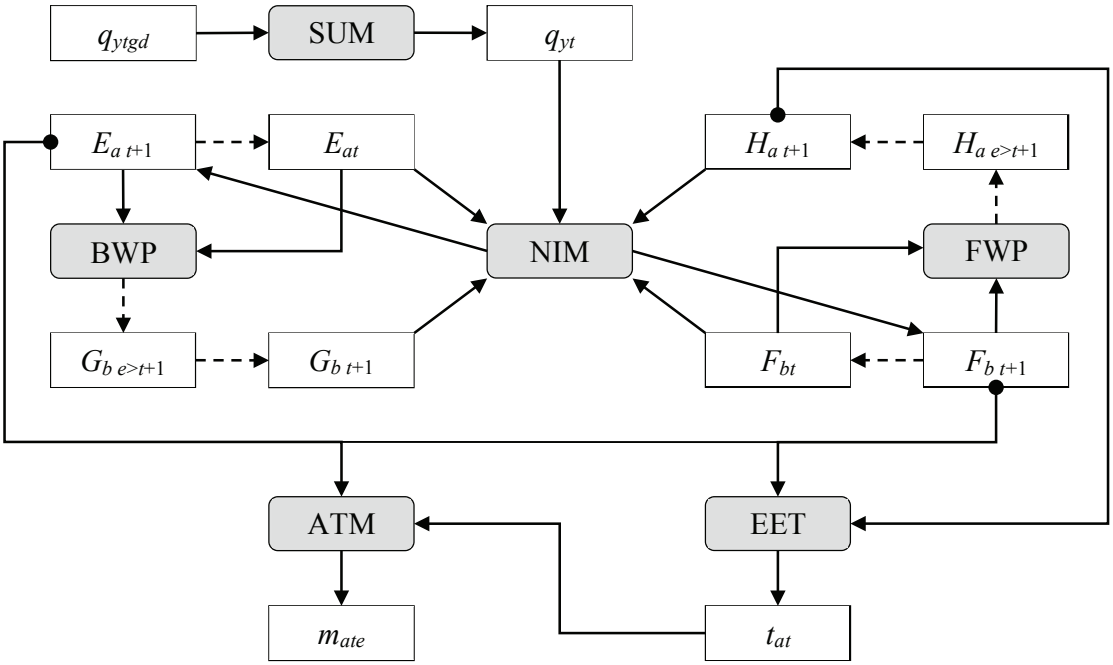

Figure 2. Schema of the simulation algorithm based on the GLTM for a given time interval. Rounded boxes are functions, while sharp boxes are variables. The dashed arrows indicate a relation between different temporal layers. The arrow with initial bullet indicate a relation that involves the entire temporal profile of a variable.

Notation of the GLTM

q_{yt}	equivalent flow (volume) passing through turn $y \in Y$ during time interval $[\tau_t , \tau_{t+1})$ with $t \in T$
F_{at}	cumulative inflow of arc $a \in A$ at time τ_t , with $t \in T$
E_{at}	cumulative outflow of arc $a \in A$ at time τ_t , with $t \in T$
G_{at}	cumulative receiving flow of arc $a \in A$ at time τ_t , with $t \in T$
H_{at}	cumulative sending flow of arc $a \in A$ at time τ_t , with $t \in T$

SUM – Turn flows specific by destination and classes are aggregated as vehicle equivalents. This allows to retrieve outflows from origins and splitting rates

NIM – Node Intersection Model. It solves the conflict between the sending flows (the maximum number of vehicles that can exit an arc up to a given instant) of its backward star and the receiving flows (the maximum number of vehicles that can enter an arc up to a given instant) of its forward star. In mergings, the downstream receiving flow is partitioned accordingly to capacities and priorities of the upstream arcs among the sending flows that are to be served. In diversions, the sending flows of all turns are served, accordingly to the FIFO rule, at a same rate which is given by the minimum ratio between the receiving and sending flow of each turn. On these bases, the outflows of the backward star and the inflows of the forward star are obtained for the current time interval.

FWP – Forward Wave Propagation. Consistently with the theory of the kinematic waves, the cumulative inflow at the beginning of the current interval is projected forward along the link to compute the cumulative sending flow at

a later time. This instant is when the shockwave between the current and previous inflow state reaches the head of the arc. The cumulative sending flow is equal to the cumulative inflow plus all the vehicles that would pass the shockwave at either two hypocritical states (it's the same). The profile point of the cumulative sending flow thus calculated is joined with the previous analogous point. The actual sending flow is given by the lower envelope of these segments.

BWP – Backward Wave Propagation. Consistently with the theory of the kinematic waves the cumulative outflow at the beginning of the current interval is projected backward along the link to compute the cumulative receiving flow at a later time. This instant is when the shockwave between the current and previous outflow state reaches the tail of the arc. The cumulative receiving flow is equal to the cumulative outflow plus all the vehicles that would pass the shockwave at either two hypercritical states (it's the same). The profile point of the cumulative receiving flow thus calculated is joined with the previous analogous point. The actual receiving flow is given by the lower envelope of these segments.

EEF – Entry-Exit Fifo. This link model yields a temporal profile of travel times for a given inflow and outflow temporal profiles, under the assumption that the First In First Out rule holds, at least on average. In this case, the travel time is the horizontal distance between the two cumulative temporal profiles of inflows and outflows (while the vertical distance is the number of vehicles on the arc). If these profiles are approximated as piecewise linear functions pivoted on a fixed temporal discretization, the approximation of this method can be as high as the duration of the time interval.

ATM – Arc Time Map. Given the cumulative temporal profile of inflows and outflows, based on the FIFO rule it is possible, not only to determine the temporal profile of travel times through the EEF model, but also to calculate the share of inflow entering during a certain interval that exits during the same or a next interval. The resulting arc time map is a crucial information for the correct propagation of flows through the network. Especially, in the presence of queue spillback from downstream, where the arc outflow is the result of the time varying exit capacity and is not simply given by the inflow scaled by the relative variation of travel time, the introduction of the arc map represents a key improvement wrt the original DUE model (Bellei et al., 2005) in terms of consistency between the supply and demand sides of the dynamic equilibrium model.

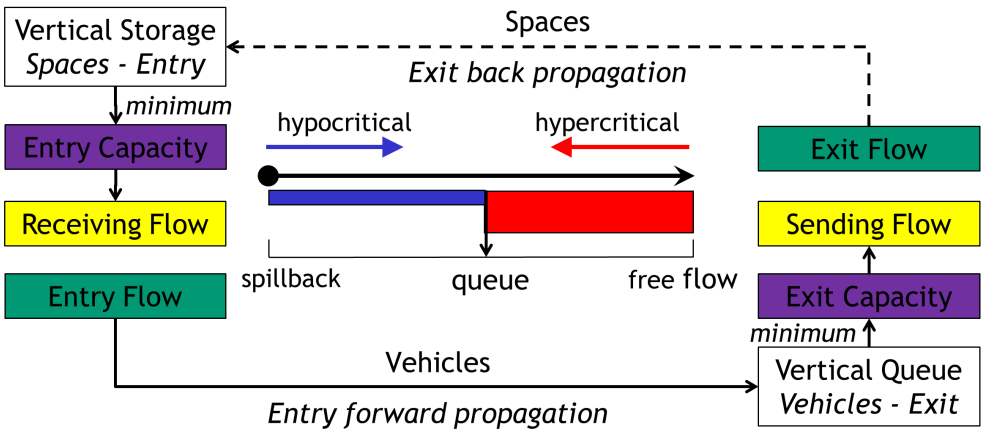

Figure 3. Link model. Forward propagation of entry flows and backward propagation of exit flows (spaces) through Newell's solution of Kinematic Wave Theory based on cumulative flows.

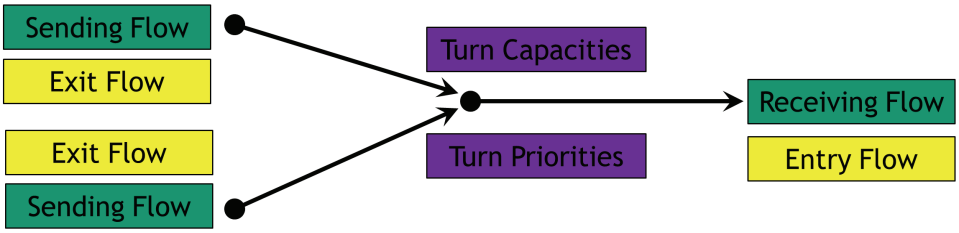

Figure 4. Node merging. Partition of scarce resource (receiving flow) among BS links, based on turn capacities and priorities. If a sending flow does not fully exploit the assigned resource the rest is shared among "hungry" arcs.

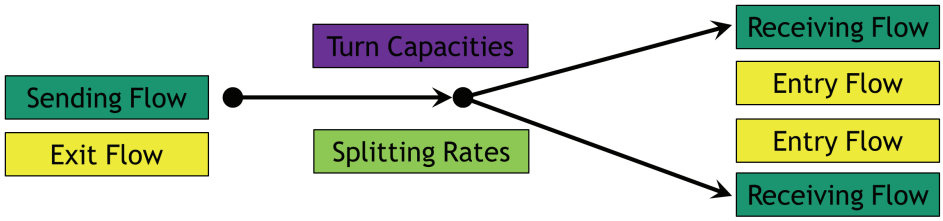

Figure 5. Node diversion. Assuming FIFO, min ratio between receiving and splitting rate by sending flow.

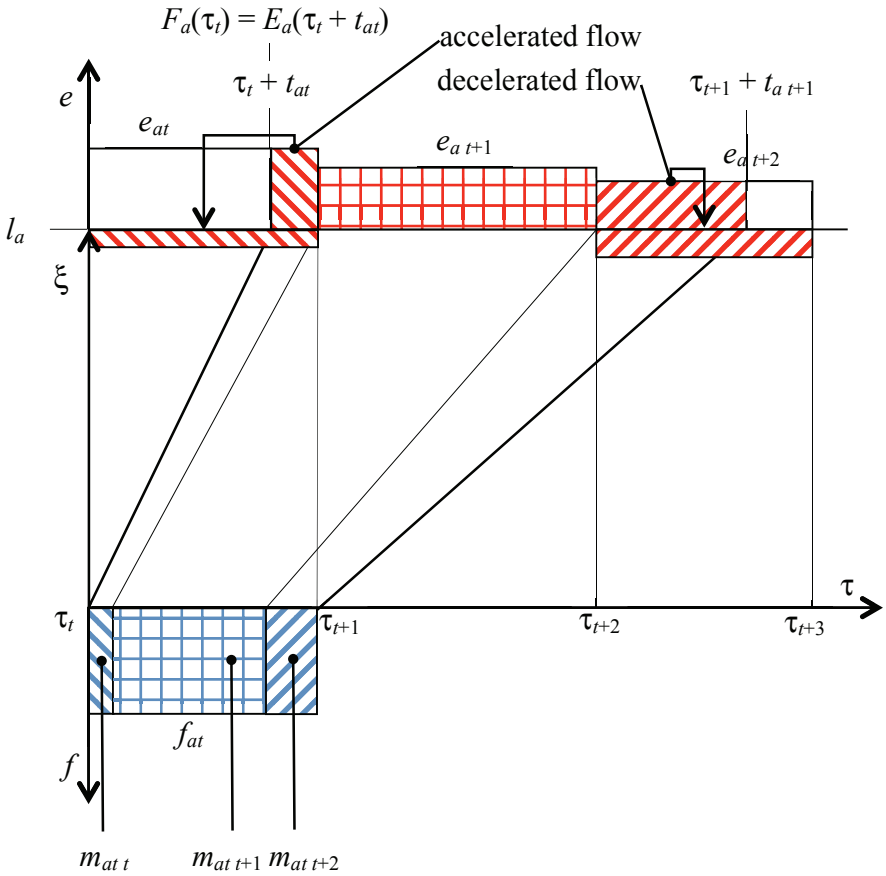

Figure 6. Flow Propagation Model and Arc Time Map.

While in Gentile et al. (2007) the consistency of the dynamic network loading between flows and travel times (including spillback and capacity constraints) is achieved only jointly with the equilibrium, in the proposed framework a proper Dynamic Network Loading based on splitting rates is obtained at each iteration. This conveys robustness to the model and allows for an easier interpretation of results, even when the level of convergence achieved is not fully satisfactory.

2.2. The Network Performance Function and the functional approach

The second approach we have considered is a revised version of the Network Performance Function (NPF) with spillback introduced in (Gentile et al., 2007), which is the base of the DUE model available in the VISUM software package. In this version, the concepts of exit and entry capacities are replaced with the NIM based on cumulative sending and receiving flows. This way, the NPF is casted in the same theoretical framework of the GLTM.

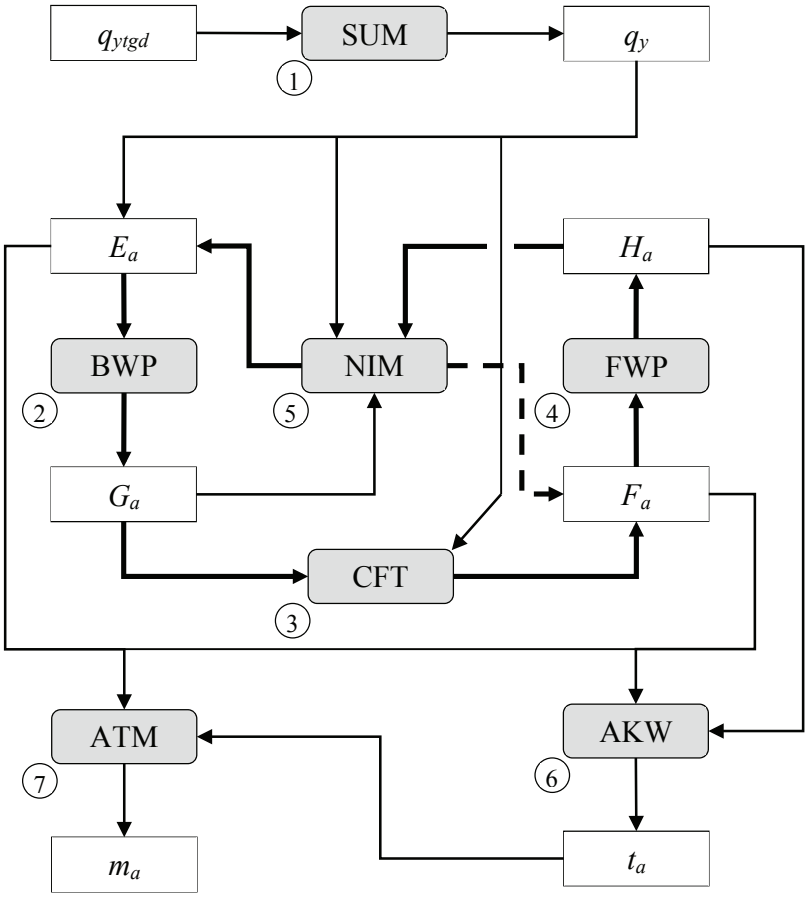

Figure 7. Schema of the NCM model with functional approach. The circled numbers indicate the sub-models sequence. The sub-sequence 2-3-4-5 forms a cycle that identifies a fixed point problem.

Thus, no new variable is introduced. Moreover, SUM, BWP, FWP and NIM are the same already described. But here each sub-model is applied separately in chronological order to the whole temporal profile, instead that using a temporal layer approach.

CFT – Capacity Flow Transformation. This sub-model spreads forward in time the cumulative inflow profile to cope with cumulative receiving flows. This is necessary to avoid multiple counting of the queue time.

AKW – Average Kinematic Waves. This link model (Gentile et al., 2005), inspired by the theory of kinematic waves, first yields a temporal profile of travel times for a given inflow temporal profile representing the hypocritical congestion, as if no queue is present at the end of the arc. Then, in case of queues, when the cumulative sending flow is higher than the cumulative outflow, the EEF model is applied. Since the typical time intervals of the NPF are long, the approximation of the EEF model in the computation of travel times can be relevant. However, in real networks sharp changes of hypercritical exit flows occur mostly in two specific occasions: the beginning and end of the queue on the arc, the beginning and end of a spillback from downstream arcs. These cases can be easily identified by comparing the cumulative sending flows vs. outflows, and the cumulative receiving flows vs. inflows, which allows introducing an additional pivot point in the piecewise linear approximation of cumulative outflows.

The sub-models BWP + CAP + FWP + NIM form a fixed point problem that can be solved in a finite number of iterations, if no gridlock occurs. Each one of these iteration is able to propagate a spillover phenomenon of one arc backward. In practice, only one iteration can be performed, leaving to the outer equilibrium iterations the task of fixing the spillback propagation as well.

Clearly, the longer the time interval the shortest the computation runtime. Therefore, if the use case does not need a second-by-second simulation, the functional approach is more flexible than the simulation approach.

A second advantage of the NPF wrt GLTM is the possibility of choosing the number of spillback (internal) iterations. Indeed, in the early outer iterations when we are far away from the equilibrium, gridlock is more likely to occur. But the resulting peaks of travel times induce a highly instable and fluctuating process to reach equilibrium. Therefore for the convergence it is better to perform only one spillback iteration per outer iteration, thus letting at early stages the queue grow more than what is physically possible on the arc where the spillover originates, instead of simulating the network necrosis produced by a gridlock (that will then disappear with further equilibrium iterations).

3. Numerical tests

3.1. Corridor to investigate Dynamic Network Loading

The corridor under analysis consist of 7 links; each one is long 200 meters. The demand is constant for the first 20 minutes and equal to 900 veh/h. The exit capacity of the 6th link is reduced from 1800 to 400 veh/h, which produces a bottleneck. The fundamental diagram of all links is plotted in Figure 8 with the ordinate axis pointing down to ensure consistency with the shockwaves of Table **Errore. L'origine riferimento non è stata trovata.**. In the simulation approach, the GLTM has intervals of one second. The time interval on the demand side are set initially to one minute. Time discretization of the supply side in the functional approach is set equal to that of the demand side. In the functional approach, exactly 7 iterations were enough to let the dynamic network loading achieve convergence, while one iteration solves the simulation approach, because no route choice is involved.

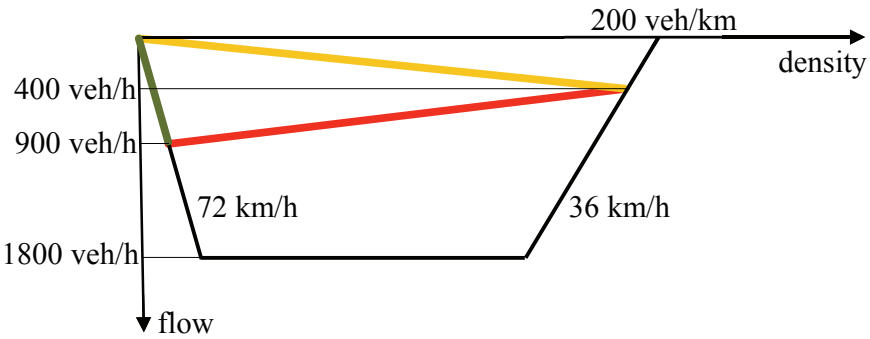

Figure 8. Trapezoidal Fundamental diagram of the corridor. Shockwaves are identified: in green the free flow; in red the queue raise; in orange the queue dissipation.

The resulting congestion pattern is relatively simple. The demand flow departed from the first link arrives after one minute at the end of the 6th link. From that instant a queue starts developing backwards along the corridor, and a shockwave (in red) separates the hypocritical states related to the demand (900 veh/h) from the hypercritical states related to the bottleneck capacity (400 veh/h). Clearly, the queue spillbacks along the link sequence. When the demand flow ends, the queue start dissipating, and a second shockwaves appears (orange). The difficulty from a modelling point of view is the correct representation of the spillback phenomenon, especially in the context of the functional approach, consistently with the propagation of waves.

Table 1. Inflows and travel times produced by the simulation approach. On the abscissa we have the timeline in seconds.

flow	0	60	120	180	240	300	360	420	480	540	600	660	720	780	840	900	960	1020	1080	1140	1200	1260	1320	1380	1440	1500	1560	1620	1680	1740	1800	1860	1920	1980	2040	2100	2160	2220	2280	2340	2400	2460	2520	2580	2640	2700	2760	2820	2880	2940	3000
1	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	150	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	150	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	600	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	900	867	400	400	400	400	400	133	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	450	900	900	900	900	900	900	900	900	900	900	900	900	750	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	300	900	900	900	900	900	900	900	900	833	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	267	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	150	900	900	900	900	717	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	133	0	0	0	0	0	0	0	0	0	0	0
7	0	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400	0	0	0	0	0	0

time	0	60	120	180	240	300	360	420	480	540	600	660	720	780	840	900	960	1020	1080	1140	1200	1260	1320	1380	1440	1500	1560	1620	1680	1740	1800	1860	1920	1980	2040	2100	2160	2220	2280	2340	2400	2460	2520	2580	2640	2700	2760	2820	2880	2940	3000
1	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
2	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	27	102	177	140	80	20	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
3	10	10	10	10	10	10	10	10	10	10	10	10	10	10	45	120	195	270	340	340	340	340	340	340	300	240	180	120	60	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
4	10	10	10	10	10	10	10	10	10	10	62	137	212	287	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	280	220	160	100	40	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
5	10	10	10	10	10	10	80	155	230	305	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	320	260	200	140	80	20	10	10	10	10	10	10	10	10	10	10	10
6	10	22	97	172	247	322	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	340	300	240	180	120	60	10	10	10	10	10	10
7	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10

c_u	0	60	120	180	240	300	360	420	480	540	600	660	720	780	840	900	960	1020	1080	1140	1200	1260	1320	1380	1440	1500	1560	1620	1680	1740	1800	1860	1920	1980	2040	2100	2160	2220	2280	2340	2400	2460	2520	2580	2640	2700	2760	2820	2880	2940	3000
1	95	145	220	294	366	440	520	595	668	741	820	895	969	1042	1120	1195	1270	1344	1420	1493	1534	1479	1419	1361	1308	1248	1188	1128	1068	1008	948	888	828	768	712	657	597	537	477	418	364	310	250	191	134	90	70	70	70	70	70

On the ordinate we have the inflow in veh/h (of the next time interval) and the travel time in sec of each link of the corridor in sequence. Thus the table provides also a dynamic representation of the resulting flow states. In green, we have hypocritical states; in yellow hypercritical states; in orange we have spillback. The last row provides the od-cost in sec for users entering at a given time.

Table 2. Difference of inflows and travel times between the simulation approach and the functional approach: in red the negative values in blue the positive ones.

flow	0	60	120	180	240	300	360	420	480	540	600	660	720	780	840	900	960	1020	1080	1140	1200	1260	1320	1380	1440	1500	1560	1620	1680	1740	1800	1860	1920	1980	2040	2100	2160	2220	2280	2340	2400	2460	2520	2580	2640	2700	2760	2820	2880	2940	3000
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	25	-25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	268	0	0	0	-168	-133	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	71	-67	-4	0	0	0	0	0	0	0	0	0	0	150	151	0	0	0	0	0	0	0	0	0	0	0	0	0	-301	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	134	-119	-15	-1	0	0	0	0	0	267	35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-35	-267	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	212	-177	-32	-3	0	301	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-168	-133	0	0	0	0	0	0	0	0	0	0
7	301	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-301	0	0	0	0	0	0

time	0	60	120	180	240	300	360	420	480	540	600	660	720	780	840	900	960	1020	1080	1140	1200	1260	1320	1380	1440	1500	1560	1620	1680	1740	1800	1860	1920	1980	2040	2100	2160	2220	2280	2340	2400	2460	2520	2580	2640	2700	2760	2820	2880	2940	3000
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-14	-45	-45	-45	-45	-10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	-29	-45	-45	-45	-45	0	0	0	0	-25	-45	-45	-45	-45	-45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	-42	-45	-45	-23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-45	-45	-45	-45	-45	-30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	1	-45	-45	-45	-45	-5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-5	-45	-45	-45	-45	-45	-10	0	0	0	0	0	0	0	0	0	0
6	0	-12	-40	-45	-45	-45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-25	-45	-45	-45	-45	-45	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 3. Inflows and travel times produced by simulation (GLTM) approach (on the left) and the functional (NPF) approach (on the right) for time intervals of 10 minutes.

flow	0	600	1200	1800	2400	3000	3600
1	900	900	0	0	0	0	0
2	885	900	15	0	0	0	0
3	870	797	133	0	0	0	0
4	855	585	360	0	0	0	0
5	813	400	400	187	0	0	0
6	587	400	400	400	13	0	0
7	360	400	400	400	240	0	0

time	0	600	1200	1800	2400	3000	3600
1	10	10	10	10	10	10	10
2	10	10	177	10	10	10	10
3	10	10	340	10	10	10	10
4	10	62	340	280	10	10	10
5	10	340	340	340	20	10	10
6	10	340	340	340	340	10	10
7	10	10	10	10	10	10	10

cost	0	600	1200	1800	2400	3000	3600
1	121	804	1301	823	378	70	70

flow	0	600	1200	1800	2400	3000	3600
1	900	900	0	0	0	0	0
2	885	900	15	0	0	0	0
3	870	829	101	0	0	0	0
4	856	624	318	2	0	0	0
5	841	412	400	147	0	0	0
6	627	400	400	371	2	0	0
7	400	400	400	400	200	0	0

time	0	600	1200	1800	2400	3000	3600
1	10	10	10	10	10	10	10
2	10	11	62	10	10	10	10
3	10	14	211	10	10	10	10
4	10	21	339	217	10	10	10
5	10	322	340	340	10	10	10
6	10	340	340	340	296	10	10
7	10	10	10	10	10	10	10

cost	0	600	1200	1800	2400	3000	3600
1	118	750	1161	763	333	70	70

The results of Table **Errore. L'origine riferimento non è stata trovata.** show that the two Network Congestion Models produce basically the same results in terms of both inflows and travel times. The only non-negligible difference is the Arc Time Map at the beginning of the simulation. Because the GLTM propagates the flows from the origin the error of the spreading effect affecting the Flow Propagation Model is not transmitted along the arc; on the contrary, in the NFP we could see some OD flow travelling at very high or very low speed because the spreading effect accelerates or delays a portion of the flow exiting from each node, as shown in Figure 6.

The main aspects of the traffic phenomenon at hand can be well captured also by considering long time intervals (10 minutes), without introducing relevant distortions, as shown in Table **Errore. L'origine riferimento non è stata trovata.**.

3.2. Fork to investigate the splitting rate issue

The network is constituted by a bottleneck with capacity of 1000 veh/h and two exiting links: North bound and South bound. A demand of 1400 veh/h is directed for the first 20 min to destination North and for the second 20 min to destination South. In the simulation approach, the first loading provides wrong splitting rates at the diversion, because actually it takes more than 20 min for all the demand directed to North to exit the bottleneck, while after 40 min a splitting rate based on capacities (that are assumed equal) is considers.

However, the GLTM computes a correct value of the bottleneck travel time, so that during the second iteration the splitting rates are adjusted and the algorithm converges.

In this case, it takes just two iterations for the dynamic network loading to adjust and provide arc flows consistent with the OD demand flows.

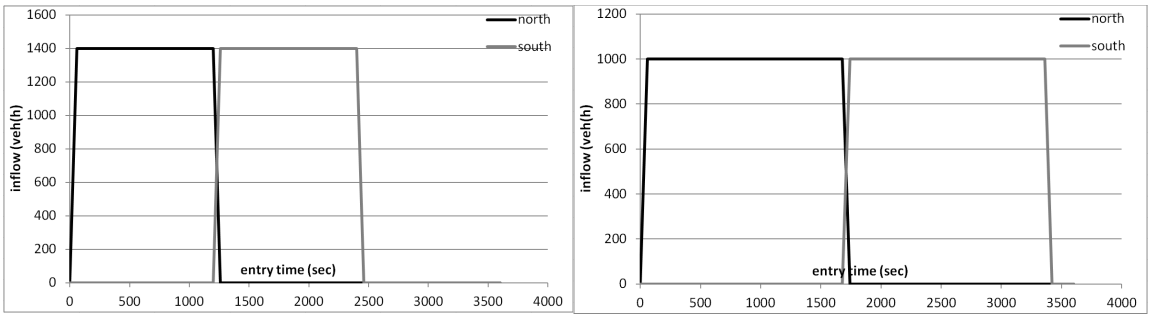

Figure 9. Inflows of the North and South bounds resulting after the first (left picture) and second (right picture) iteration.

3.3. Dipole to investigate Dynamic User Equilibrium

Dipole is a simple network with two alternative roads to go from A to B: the main way and a deviation. The main way is 1.1 km long and has a bottleneck at the end with a capacity of 500 veh/h. The deviation is 10 km long. The capacity of all links is 2000 veh/h, the speed is 60 veh/h, the jam wave speed is 25 km/h, the jam density is 200 veh/km. The fundamental diagram of the deviation has a quadratic hypocritical shape, which implies congestion along the link due to vehicle interaction (desired speed variance and consequent difficulties in overtaking). The travel demand is constant for the first 40 minutes and equal to 1500 veh/h. Time discretization is identical to the corridor case.

Initially, the bottleneck is definitely the convenient alternative; but as the queue grows, its travel time increases, up to the point when the two alternatives are equivalent. At that point a share of demand flows starts choosing the deviation, instead of bottleneck. This share increases progressively, as the deviation gets also congested, although in hypocritical states, while the queue still grows at a lesser rate. At equilibrium, clearly the amount of demand flow choosing the bottleneck is exactly equal to the bottleneck capacity. Finally, the demand stops and the queue starts to dissipate.

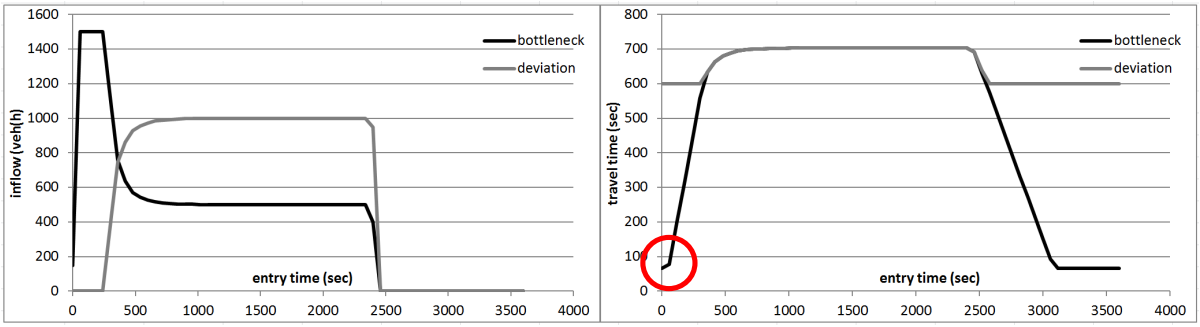

Figure 10. Inflows and travel times produced by the simulation approach (**GLTM**) for time intervals of **1 minute**.

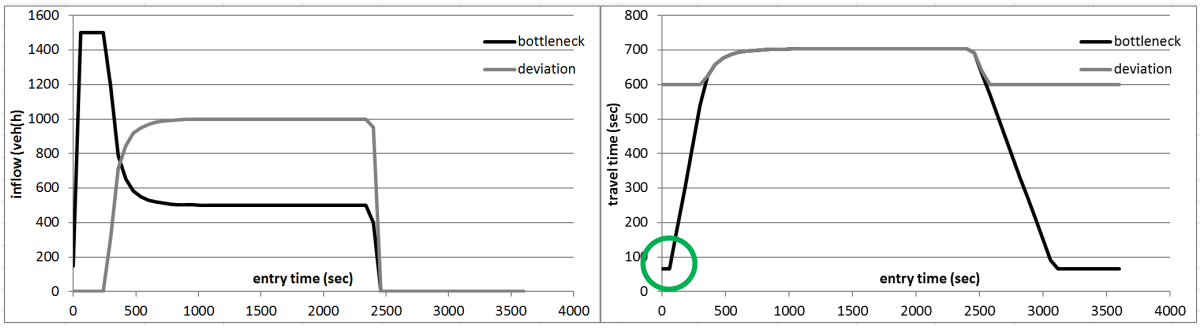

Figure 11. Inflows and travel times produced by the functional approach (**NPF**) for time intervals of **1 minute**.

The only (minor) differences that can be perceived between the two proposed models are relate to the travel time computation in absence of flow, where the current NPF implementation is hybridized with the AKW model. Indeed. travel times are not a direct output of the GLTM, and they have to be retrieved from the entry-exit model, which is affected by approximations.

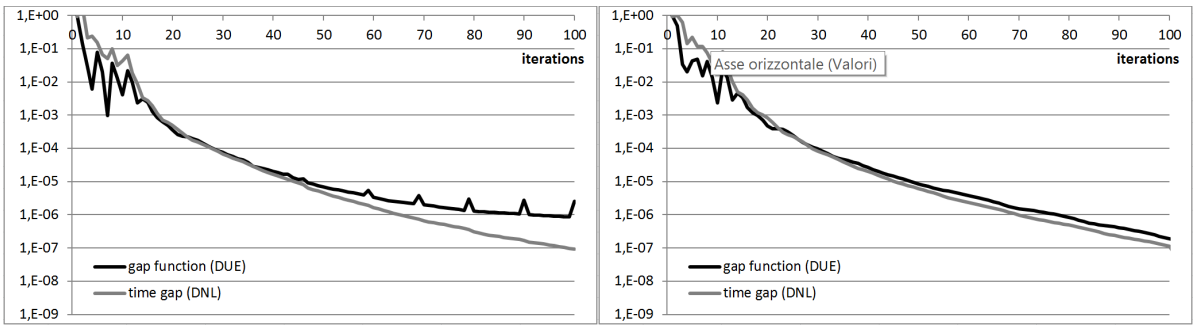

Figure 12. Convergence pattern to the equilibrium with simulation (**GLTM**) approach (on the left) and with the functional (**NPF**) approach (on the right) for time intervals of **1 minute**.

As shown in Figure 12, the Gradient Projection algorithm is able to solve both models at any level of precision. However, the functional approach shows better performances in this respect.

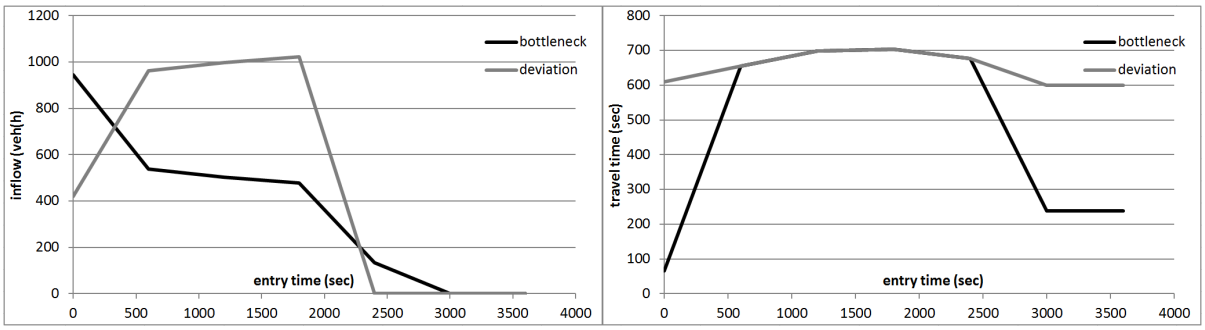

Figure 13. Inflows and travel times produced the simulation approach (**GLTM**) for time intervals of **10 minutes**.

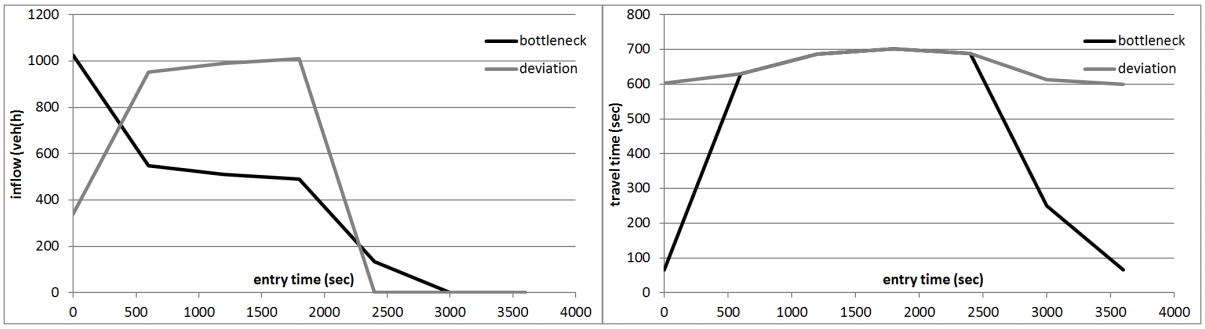

Figure 14. Inflows and travel times produced by the functional approach (**NPF**)for time intervals of **10 minutes**.

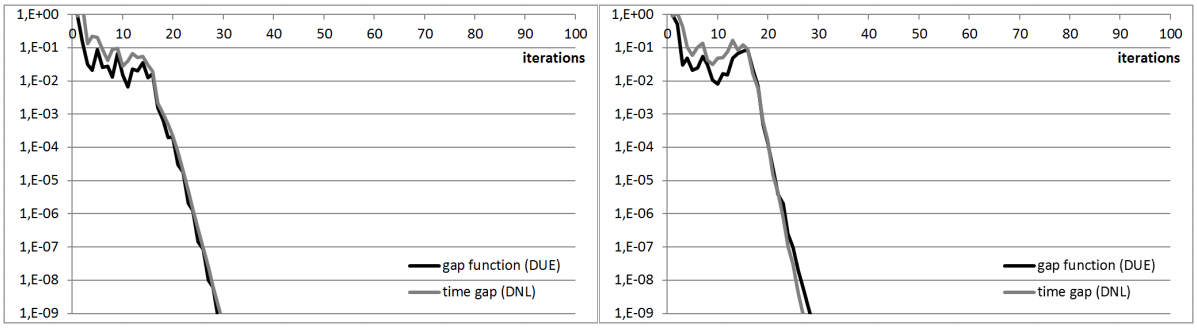

Figure 15. Convergence pattern to the equilibrium with simulation (**GLTM**) approach (on the left) and with the functional (**NPF**) approach (on the right) for time intervals of **10 minutes**.

It is interesting to notice that, considering time intervals of 10 minutes, we still have a realistic picture of what happens on the network, despite the approximations induced by the rough time discretization. Convergence is much easier: after all the number of variables that we are moving is 10 times less.

4. Discussion and conclusions

In this paper a general framework for modelling DTA using a Link Transmission Model has been presented. A more classical approach, where the GLTM is used as a simulation tool to reproduce network congestion on the supply side, is compared with a more innovative approach, where the same components of the GLTM are used to provide a Network Performance Function that is a relation among temporal profiles. The outputs of the two resulting models in terms of flows and travel times are basically identical, regardless the temporal discretization that is used

on the demand side for the route choice and the flow propagation. This gives a great flexibility in the application of the two approaches in different use cases which constitutes the main achievement of this work.

From a computational perspective, the simulation approach is more consuming in terms of runtime and memory because the internal temporal layers must be anyhow rather brief (few seconds), so as to satisfy the constraint that no kinematic wave, propagating hypocritical flow states forward from the arc tail or hypercritical states backward from its head, is faster than the ratio between link length and interval duration.

On the contrary, the functional approach allows for long time intervals (few minutes). This advantage comes with a price: in presence of spatial non-separability, i.e. node stream conflicts and spillback, the dynamic network loading model requires some internal iterations to propagate the effects of queue congestion upstream. This inconvenient is less relevant in the context of an equilibrium problem where it is possible to perform only one internal iteration, leaving to the outer iterations the task of adjusting also the dynamic network loading, beside the route choice.

The simulation approach, instead, implies a straightforward computation in chronological order without internal iterations. However, the splitting rates (turning fractions) used in this process are consistent only with the flow propagation at the demand side, while OD flows are not satisfied on the supply side at the generic iteration; this inconsistency is solved only at equilibrium. For this reason, after stopping the equilibrium process (sometimes this happens in practice because the maximum number of iterations is reached, without satisfying the desired convergence criterion), it is convenient to perform a limited number of additional iterations with fixed route choices (keeping the same temporal profiles of the arc conditional probabilities by destination) to adjust the dynamic network loading and thus achieving OD flow consistency also on the supply side; this process indeed converges much faster than equilibrium and a few iterations (typically less than 10) are enough.

More specifically, the complexity of one outer iteration of equilibrium is: $\alpha \cdot |Z| \cdot |A| \cdot |T| + \text{NCM}$;
the complexity of the NCM with the simulation approach is: $\beta \cdot |I| \cdot |A|$;
the complexity of the NCM with the functional approach is: $\beta \cdot n \cdot |T| \cdot |A|$;
where $|Z|$ is the number of zones (destinations), $|A|$ is the number of arcs, $|T|$ is the number of time intervals on the demand side, $|I|$ is the number of simulation intervals, n is the number of internal iterations (can be one), α and β are the number of cpu operations to process each network element respectively on the demand and supply side.

In real applications we can consider: $|I| \approx 100 \cdot |T|$ and $n = 1$. However, in the context of an equilibrium problem the higher complexity of the simulation approach does not affect substantially the overall complexity, because typically α and β have the same order of magnitude, while $|Z| >> 100$.

Two use cases are discussed in the following: real-time traffic management, off-line transport planning. In the first case, run times are a relevant issue, but precise convergence to equilibrium is not compelling (under some circumstances, we can even assume fixed splitting rates), while consistency in the dynamic network loading is highly desirable. Moreover, if signal setting and coordination are an issue, a higher detail and granularity on the supply side is required. This induces to adopt the simulation approach.

In the second case, convergence to equilibrium is compelling because it allows a consistent comparison of design scenarios, but there are softer constraints in terms of computation runtime. Here, we are interested in the evaluation of Key Performance Indicators resulting from the assignment, while a more detailed representation of the supply side can be achieved as a post process, after equilibrium is reached. This induces to adopt the functional approach, which proves to be more robust in terms of convergence (for example, fictitious gridlock states that would appear during the equilibration process can be avoided considering $n = 1$), and maybe a GLTM with the splitting rates resulting from equilibrium on top of it.

Other numerical test not presented here for brevity prove the validity of the these conclusions on larger networks. However, in this case convergence is not as fast as for the simple dipole, due to the higher overall complexity of the problem. In general, with 100 iterations of Gradient Projection we can expect to reach 10^{-4}, while with 100 iterations of MSA we cannot expect to reach more than 10^{-3}. These one or two additional levels of convergence in logarithmic scale imply 10 to 100 times higher precision, which can mean a lot in terms of solution quality and robustness. A more detailed analysis of the Gradient Projection algorithm compared with the MSA has been addressed with respect to larger networks in Gentile (2014b).

Table 4. Strengths and weakness of the two proposed methods: simulation (GLTM), functional (NPF).

Feature	GLTM	NPF
Possibility of using long time intervals with relevant runtime reductions	no	yes
Possibility of choosing the number of internal spillback iterations to avoid gridlock at early external equilibrium iterations	no	yes
Full precision of the spillback calculation	yes	no
Spreading effect that accelerates or delays flows confined to a local arc distortion	yes	no
Supply side flows consistent with OD demand at each iteration and not only at equilibrium	no	yes
Travel time computation hybridized with the AKW model for hypocritical states	no	yes
Smoother convergence to equilibrium	no	yes

References

Adamo V., Astarita V., Florian M., Mahut M., Wu J., 1999. Modeling the spillback of congestion in link based dynamic network loading models: A simulation model with application. In Transportation and Traffic Theory, ed. A. Ceder, Pergamon-Elsevier, New York, USA, pp. 555-573.

Barcelo J., 2010. Fundamentals of Traffic Simulation, Springer, New York, USA,

Bellei G., Gentile G., Papola N., 2005. A within-day dynamic traffic assignment model for urban road networks. Transportation Research B 39, pp. 1-29.

Ben-Akiva M., Bierlaire M., Bottom J., Koutsopoulos N., Mishalani R., 1997. Development of a route guidance generation system for real-time application. Proceedings of the 8th IFAC symposium on transportation systems, Chania, Greece.

Ben-Akiva M, Bierlaire M, Koutsopoulos H., Mishalani R., 2002. Real-time simulation of traffic demand–supply interactions within DynaMIT. In Transportation and network analysis: current trends. Miscellanea in honour of Michael Florian, ed.s M. Gendreau and P. Marcotte, Kluwer, Boston, USA, pp. 19-36.

Chabini I., 1998. Discrete dynamic shortest path problems in transportation applications: complexity and algorithms with optimal run time. Transportation Research Record 1645, pp. 170-175.

Daganzo C., 1994. The cell transmission model: A simple dynamic representation of highway traffic. Transportation Research B 28, pp. 269-287.

Daganzo, C. 1995 The cell transmission model, Part II: Network traffic. Transportation Research B 29, pp. 79-93.

Dial R.B., 2006. A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration. Transportation Research B 40, pp.917-936.

Friesz T., Bernstein D., Smith T., Tobin R., Wie B., 1993. A variational inequality formulation of the dynamic network user equilibrium problem. Operations Research 41, pp. 179-191.

Gentile G., 2010a. Traffic assignment: can dynamic models offer more?. In Proceedings of Networks for Mobility, 5th International Symposium, ed.s U. Martin et al., FOVUS, Stuttgart, Germany, pp. 1-5.

Gentile G., 2010b. The General Link Transmission Model for Dynamic Network Loading and a comparison with the DUE algorithm. In New developments in transport planning: advances in Dynamic Traffic Assignment (selected papers from the DTA 2008 Conference, Leuven), ed.s L.G.H. Immers, C.M.J. Tampere, F. Viti, Transport Economics, Management and Policy Series, Edward Elgar Publishing, MA, USA, pp. 153-178.

Gentile G., Meschini L., Noekel K., 2006. DUE: Dynamic User Equilibrium. Technical Report in VISUM 10 Manual, Karlsruhe, Germany.

Gentile G., 2014. Local User Cost Equilibrium: a bush-based algorithm for traffic assignment. Transportmetrica A: Transport Science 10, 15-54.

Gentile G., 2014b. Solving a Dynamic Traffic Assignment model with implicit path enumeration using Gradient Projection algorithms. Proceedings of the 5th International Symposium on Dynamic Traffic Assignment - DTA 2014, Salerno, Italy.

Gentile G., Meschini L., 2004. Fast heuristics for continuous dynamic shortest paths and all-or-nothing assignment. Proceedings of the AIRO 2004, Lecce, Italy.

Gentile G., Meschini L., Papola N., 2005. Macroscopic arc performance models with capacity constraints for within-day dynamic traffic assignment. Transportation Research B 39, pp. 319-338.

Gentile G., Meschini L., Papola N., 2007. Spillback congestion in dynamic traffic assignment: a macroscopic flow model with time-varying bottlenecks. Transportation Research B 41, pp. 1114-1138.

Gentile G., Meschini L., 2011. Using dynamic assignment models for real-time traffic forecast on large urban networks. In Proceedings of the 2nd International Conference on Models and Technologies for Intelligent Transportation Systems, Leuven, Belgium.

Jayakrisham R., Mahmassani H., Hu T., 1994. An evaluation tool for advanced traffic information and management systems in urban networks. Transportation Research C 2, pp. 129-147.

Heydecker B., Addison J., 1998. Traffic models for dynamic traffic assignment. In Transport networks: recent methodological advances, ed M.G.H. Bell, Pergamon-Elsevier, Oxford, UK, pp. 35-49.

Himpe W., Corthout R., Tampère C, 2013. An Implicit Solution Scheme for the Link Transmission Model. Proceeding of the 16th IEEE ITSC, The Hague, The Netherlands.

Lo H., Szeto W., 2002. A cell-based variational inequality formulation of the dynamic user optimal assignment problem. Transportation Research B 36, pp. 421-443.

Mahmassani H., 2001. Dynamic network traffic assignment and simulation methodology for advanced system management applications. Network and Spatial Economics 1, pp. 267.292.

Mahut M., Florian M., Tremblay N., 2008. Comparison of assignment methods for simulation based dynamic-equilibrium traffic assignment. Presented at the 87th Annual Meeting of the Transportation Research Board, Washington DC, USA.

Newell G.,1993. A simplified theory of kinematic waves in highway traffic, part I: general theory; part II: queuing at freeway bottlenecks; part III: multi-destination flows. Transportation Research B 27, pp. 281-313.

Pallottino S., Scutellà M.G., 1998. Shortest path algorithms in transportation models: classical and innovative aspects. Equilibrium and advanced transportation modelling, ed.s P. Marcotte, S. Nguyen, Kluwer Academic Publisher, Dordrecht, The Netherlands, pp. 245-281.

Papageourgiou M., 1990. Dynamic modelling, assignment and route guidance in traffic networks. Transportation Research B 24, pp. 471-95.

Ramadurai G., Ukkusuri S., 2011. B-Dynamic: An efficient algorithm for Dynamic User Equilibrium assignment in activity-travel networks. Computer-Aided Civil and Infrastructure Engineering 26, pp. 254-269.

Ran B., Boyce D., 1994. Dynamic urban transportation network models: theory and implications for intelligent vehicle-highway systems. Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, New York, USA.

Smith M., Mounce R., 2011. A splitting rate model of traffic re-routeing and traffic control. Transportation Research B 45, pp. 1389-1409.

Yperman I. (2007) The Link Transmission Model for Dynamic Network Loading. PhD Thesis, Katholieke Universiteit Leuven.